

# Improvement of Monte Carlo estimates with covariance-optimized finite-size scaling at fixed phenomenological coupling

Francesco Parisen Toldin

Max-Planck-Institut für Physik Komplexer Systeme, Nöthnitzer Strasse 38, D-01187 Dresden\*

In the finite-size scaling analysis of Monte Carlo data, instead of computing the observables at fixed Hamiltonian parameters, one may choose to keep a renormalization-group invariant quantity, also called phenomenological coupling, fixed at a given value. Within this scheme of finite-size scaling, we exploit the statistical covariance between the observables in a Monte Carlo simulation in order to reduce the statistical errors of the quantities involved in the computation of the critical exponents. This method is general and does not require additional computational time. This approach is demonstrated in the Ising model in two and three dimensions, where large gain factors in CPU time are obtained.

PACS numbers: 05.10.Ln, 64.60.an, 05.70.Jk, 05.50.+q

The Monte Carlo (MC) method is a powerful numerical technique which allows to study models with a large number of degrees of freedom [1, 2]. One of the main features of the method is its flexibility: arbitrary statistical systems can be investigated in a regime where other methods (e.g., perturbative calculations) may not be available or efficient. For this reason it is one of the most used tools in critical phenomena [3]. MC integration allows to compute quantities with, at least in principle, arbitrary precision. The accuracy of the results is proportional to the inverse square root of the computational time, hence even a small improvement in the statistical error bars can represent a significant gain in terms of computational resources. Several strategies have been developed in order to improve the accuracy: one of the most common methods is the use of improved estimators, which have the same expectation value of the desired quantities, but reduced error bars (see, e.g., Ref. [2]). More recently, it has been proposed to use a covariance analysis in order to compute the optimal weighted average of different estimates of a critical exponent [4] and to reduce the statistical error of an observable by adding a control variate, whose mean value vanishes [5].

In the context of numerical investigations of critical phenomena, finite-size scaling (FSS) techniques play an important role: they allow to extract the critical properties of a model from the data obtained in a region of parameters where the correlation length  $\xi$  is of the order of the linear size  $L$  of the system. In the FSS analysis of MC data, instead of computing the observables at fixed Hamiltonian parameters, one can choose to fix a renormalization-group (RG) invariant quantity  $R(\beta, L)$ , also called phenomenological coupling, at a given value  $R_f$ . In this method, introduced in Ref. [6] and discussed also in Refs. [7, 8], one determines for every  $L$  an inverse temperature  $\beta_f(L)$  such that  $R(\beta = \beta_f(L), L) = R_f$ . Then all the observables are computed at  $\beta = \beta_f(L)$  and the statistical errors can be estimated using standard Jackknife techniques (see, e.g., Ref. [9]). It is easy to show that  $\beta_f(L \rightarrow \infty) \rightarrow \beta_c$ , with  $\beta_c$  the inverse critical temperature. It has been observed that, thanks to cross correlations, this method gives in some cases reduced error bars [7, 8]. Error

reduction has been reported also in the related FSS method of phenomenological renormalization, where two system sizes are enforced to share a common value of  $R = \xi/L$  [10].

In this work we study the improvement of the accuracy of MC estimates by considering FSS at fixed phenomenological coupling, to be chosen as a linear combination of a given set of phenomenological couplings. By exploiting the covariance between the observables, we show how to choose the combination which minimizes the statistical errors of the quantities involved in the computation of the critical exponents. We illustrate this method by considering the Ising model in two and three dimensions simulated using, separately, Metropolis and Wolff single-cluster algorithms.

**Finite-size scaling at fixed phenomenological coupling**  
A MC simulation consists of a Markov chain whose equilibrium distribution is the Boltzmann-Gibbs measure [1]. Any observable is obtained using a statistical estimator, i.e., a random variable whose expectation value is the thermal average of the observable. Since all the observables are calculated from the same configurations generated in the simulation, they are in general correlated with each other. We denote a random variable with a hat  $\hat{A}$ , its expectation value by  $E[\hat{A}] = A$ , and its fluctuation around the expectation value by  $\delta\hat{A} \equiv \hat{A} - A$ . In what follows, the dependence of the observables on the size of the system is implicit. Let us consider a generic RG-invariant quantity  $R$ , which is sampled by an estimator  $\hat{R}$ , and that we fix to the value  $R_f$ :

$$\hat{R}(\beta_f) = R_f. \quad (1)$$

The actual computation of an estimate of  $R$ , and hence of  $\beta_f$ , involves an average over the entire MC run, while the statistical error bars resulting from the fluctuations  $\delta\hat{R}$ ,  $\delta\beta_f$  can be estimated using the Jackknife technique. This method consists of dividing the set of MC data into  $N_{\text{bin}}$  consecutive blocks of equal size, which has to be much larger than the autocorrelation time, and then constructing  $N_{\text{bin}}$  Jackknife blocks by taking the set of all data except one block. From the data of every single Jackknife block, one computes an estimate of  $R$  and of  $\beta_f$  by solving Eq. (1), as well as an estimate of any other observable  $O$  at the inverse temperature  $\beta_f$ , as determined from the same Jackknife block. This procedure results in  $N_{\text{bin}}$  (cor-

\* parisen@pks.mpg.de

related) estimators  $\widehat{O}_f^{(j)}$  of the observable  $O$  at fixed  $R = R_f$ , from which the statistical error bars are computed using the standard Jackknife estimator [9]:

$$\widehat{\sigma}^2(O_f) = \frac{N_{\text{bin}} - 1}{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} \left( \widehat{O}_f^{(j)} - \widehat{O}_f \right)^2, \quad (2)$$

where  $\widehat{O}_f$  is an estimator of  $O$  at fixed  $R = R_f$ . Note that  $\widehat{\sigma}^2(O_f)$  is itself a random variable, whose expectation value is the sought variance. Solving Eq. (1) requires an extrapolation of the MC data sampled at  $\beta_{\text{run}}$  to  $\beta = \widehat{\beta}_f$ , which can be done by using reweighting techniques or by computing a Taylor expansion of the observables on  $\beta$ . This in practice requires  $\widehat{\beta}_f \simeq \beta_{\text{run}}$ , or equivalently  $\widehat{R} \simeq R_f$ . Employing a first-order Taylor expansion of  $\widehat{R}(\widehat{\beta}_f)$  around  $\beta_{\text{run}}$ ,  $\widehat{\beta}_f$  is obtained as

$$\widehat{\beta}_f \simeq \beta_{\text{run}} - \frac{\widehat{R} - R_f}{\widehat{R}'}, \quad (3)$$

where  $\widehat{R}'$  is an estimator of the derivative of  $R$  with respect to  $\beta$ . All the other observables are then computed at the inverse temperature  $\widehat{\beta}_f$ . From an observable  $O$  sampled by the estimator  $\widehat{O}$  we obtain an observable  $O_f$ , which is sampled by the random variable  $\widehat{O}_f$ . Using Eq. (3),  $\widehat{O}_f$  is given by

$$\widehat{O}_f \simeq \widehat{O} - \widehat{O}' \frac{\widehat{R} - R_f}{\widehat{R}'}, \quad (4)$$

where  $\widehat{O}'$  is an estimator of the derivative of  $O$  with respect to  $\beta$  and as before we have used a first-order Taylor expansion of  $\widehat{O}$  around  $\beta_{\text{run}}$ . Expanding Eq. (4) around the expectation value we obtain, to the first order in the fluctuations,

$$\widehat{O}_f \simeq \widehat{O} - \frac{O'}{R'} \widehat{\delta R} - (R - R_f) \left( \frac{\widehat{\delta O}'}{R'} - \frac{O'}{R'^2} \widehat{\delta R}' \right). \quad (5)$$

The second term in Eq. (5) represents the modification in the fluctuation of the observable  $O$  which is due to the fact that we have fixed  $R$ , while the third term is related to the variation in the expectation value due to the extrapolation of the MC data to  $\beta = \beta_f$ . Since, as mentioned above,  $R \simeq R_f$ , this last term is a correction, which can be neglected in first approximation. In the presence of  $N$  RG-invariant quantities  $\{R_i\}$ ,  $i = 1, \dots, N$ , we can apply the above method by keeping fixed a generic linear combination of  $\{R_i\}$ ,  $R = \sum_i \lambda_i R_i$ . In this case generalizing Eq. (5) and neglecting its last term, we obtain

$$\widehat{O}_f(\{\lambda_i\}) = \widehat{O} - \frac{O' \sum_{i=1}^N \lambda_i \widehat{\delta R}_i}{\sum_{i=1}^N \lambda_i R'_i}. \quad (6)$$

Now we consider the following problem: find the coefficients  $\lambda_i$  which minimize the variance of  $\widehat{O}_f(\{\lambda_i\})$ . Notice that Eq. (6) is invariant under a global rescaling of the coefficients  $\lambda_i \rightarrow \mu \lambda_i$ ,  $\mu \neq 0$ : this simply reflects the equivalence of

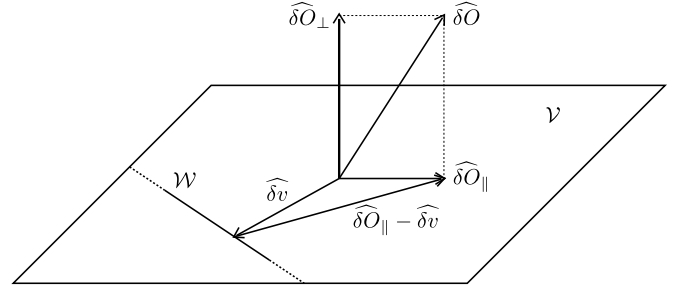


FIG. 1. Decomposition of the vector  $\widehat{\delta O}$  into two components orthogonal  $\widehat{\delta O}_\perp$  and parallel  $\widehat{\delta O}_\parallel$  to the space  $\mathcal{V} = \text{Span} \langle \delta R_1, \dots, \delta R_N \rangle$ . The affine subspace  $\mathcal{W}$  is given by Eq. (7) which represents the normalization on the parameters  $\{\lambda_i\}$ . The minimum length for  $\widehat{\delta O}_\parallel - \widehat{\delta v}$  is obtained when it is orthogonal to  $\mathcal{W}$ .

fixing  $R$  or  $\mu R$ . It is convenient to fix the normalization by requiring that

$$\sum_{i=1}^N \lambda_i R'_i = O', \quad (7)$$

so that we have to minimize the variance of

$$\widehat{\delta O}_{f,\text{norm}}(\{\lambda_i\}) = \widehat{\delta O} - \sum_{i=1}^N \lambda_i \widehat{\delta R}_i \quad (8)$$

with the constraint of Eq. (7). The subscript norm in Eq. (8) reminds that  $\widehat{\delta O}_{f,\text{norm}}$  is chosen with the normalization given in Eq. (7). The optimal  $\{\lambda_i\}$  parameters can be found using the method of Lagrangian multipliers. Here, we point out a geometrical interpretation of the problem which allows to find the solution immediately. The space of random variables with null expectation value is a real vector space. A positive-definite scalar product  $\langle, \rangle$  between two vectors  $\widehat{\delta A}$  and  $\widehat{\delta B}$  can be defined by the covariance  $\text{COV}[\widehat{\delta A}, \widehat{\delta B}]$

$$\langle \widehat{\delta A}, \widehat{\delta B} \rangle \equiv \text{COV}[\widehat{\delta A}, \widehat{\delta B}] = E[\widehat{\delta A} \widehat{\delta B}] - E[\widehat{\delta A}] E[\widehat{\delta B}], \quad (9)$$

The square length of a vector corresponds to the variance of a random variable. Let  $\mathcal{V} = \text{Span} \langle \widehat{\delta R}_1, \dots, \widehat{\delta R}_N \rangle$  be the  $N$ -dimensional subspace generated by  $\{\widehat{\delta R}_i\}$ . According to Eq. (8),  $\widehat{\delta O}_{f,\text{norm}}$  is obtained by subtracting to  $\widehat{\delta O}$  a vector  $\widehat{\delta v} = \sum_i \lambda_i \widehat{\delta R}_i \in \mathcal{V}$ , whose components in the (nonorthogonal) base  $\{\widehat{\delta R}_i\}$  are given by the coefficients  $\{\lambda_i\}$ . By decomposing  $\widehat{\delta O}$  into two components orthogonal  $\widehat{\delta O}_\perp$  and parallel  $\widehat{\delta O}_\parallel$  to  $\mathcal{V}$ , one finds that the variance of  $\widehat{\delta O}_{f,\text{norm}}$  is given by

$$\text{VAR}[\widehat{\delta O}_{f,\text{norm}}] = \text{VAR}[\widehat{\delta O}_\perp] + \text{VAR}[\widehat{\delta O}_\parallel - \widehat{\delta v}]. \quad (10)$$

Furthermore, the constraint of Eq. (7) represents an affine subspace  $\mathcal{W}$  of dimension  $N - 1$ , to which  $\widehat{\delta v}$  belongs. Minimizing the variance of  $\widehat{\delta O}_{f,\text{norm}}$  consists in finding a vector  $\widehat{\delta v} \in \mathcal{W}$  such that the length of  $\widehat{\delta O}_\parallel - \widehat{\delta v}$  is minimum.

$L$	$\chi$	$(dU_4/d\beta)/L^d$	$(dU_6/d\beta)/L^d$	$(dR_\xi/d\beta)/L^d$
Metropolis dynamics				
<b>8</b>	41.153(99)	-0.06617(99)	-0.1881(30)	0.18353(90)
imp	40.878(29)	-0.06490(17)	-0.18935(44)	0.18074(37)
gain	11.5	35.4	47.6	6.0
<b>16</b>	140.60(60)	-0.03193(93)	-0.0909(28)	0.09045(73)
imp	138.90(14)	-0.03285(14)	-0.09475(38)	0.08836(34)
gain	18.4	43.6	54.2	4.7
<b>32</b>	470.8(2.6)	-0.01685(61)	-0.0482(19)	0.04490(42)
imp	469.71(49)	-0.016369(83)	-0.04714(23)	0.04409(20)
gain	27.2	54.9	68.2	4.5
<b>64</b>	1578.5(5.8)	-0.00812(20)	-0.02334(61)	0.02151(16)
imp	1578.5(1.6)	-0.008121(42)	-0.02336(12)	0.02149(10)
gain	13.3	22.9	27.5	2.6
<b>128</b>	5284(28)	-0.00433(15)	-0.01251(47)	0.01097(11)
imp	5316.8(6.0)	-0.004078(24)	-0.011736(67)	0.010956(62)
gain	22.1	39.8	48.9	2.9
Wolff dynamics				
<b>8</b>	41.431(44)	-0.06448(45)	-0.1828(13)	0.18634(60)
imp	40.829(18)	-0.06362(18)	-0.18572(46)	0.18138(38)
gain	5.9	6.3	8.5	2.5
<b>16</b>	139.70(23)	-0.03294(37)	-0.0942(11)	0.08962(45)
imp	138.000(98)	-0.03237(13)	-0.09353(35)	0.08848(31)
gain	5.4	8.4	9.6	2.2
<b>32</b>	470.32(84)	-0.01616(21)	-0.04632(63)	0.04359(27)
imp	469.01(35)	-0.016312(77)	-0.04691(21)	0.04362(17)
gain	5.9	7.6	8.7	2.4
<b>64</b>	1582.8(3.2)	-0.00820(11)	-0.02353(34)	0.02179(15)
imp	1581.4(1.2)	-0.008177(48)	-0.02350(13)	0.02175(10)
gain	6.5	5.7	6.7	2.1
<b>128</b>	5327(12)	-0.004102(70)	-0.01176(21)	0.010983(89)
imp	5321.2(4.3)	-0.004075(29)	-0.011696(81)	0.010883(63)
gain	7.7	5.6	6.5	2.0

TABLE I. Results for the two-dimensional (2D) Ising model with Metropolis and Wolff dynamics, separately. For each lattice size  $L$ , we compare standard analysis (first line), with the method explained in this work (imp). The approximate gain factor in CPU time is denoted with gain.

This situation is illustrated in Fig. 1. The minimum length for  $\widehat{\delta O}_{\parallel} - \widehat{\delta v}$  is then obtained by choosing  $\widehat{\delta v}$  such that  $\widehat{\delta O}_{\parallel} - \widehat{\delta v}$  is orthogonal to  $\mathcal{W}$ . With this insight, the optimal coefficients  $\{\lambda_i\}$  are calculated as

$$\lambda_i = -\frac{\mathbf{R}'^T \mathbf{M}^{-1} \mathbf{N} - O'}{\mathbf{R}'^T \mathbf{M}^{-1} \mathbf{R}'} (\mathbf{M}^{-1} \mathbf{R}')_i + (\mathbf{M}^{-1} \mathbf{N})_i, \quad (11)$$

where the matrix  $\mathbf{M}$  and the vector  $\mathbf{N}$  are defined as

$$\mathbf{M}_{ij} \equiv \langle \widehat{\delta R}_i, \widehat{\delta R}_j \rangle, \quad (12)$$

$$\mathbf{N}_i \equiv \langle \widehat{\delta O}, \widehat{\delta R}_i \rangle, \quad (13)$$

and we have introduced the vector  $\mathbf{R}'_i \equiv R'_i$ . The quantities involved in Eq. (11) can be estimated from MC data by using standard Jackknife techniques [9]. It is useful to observe that

the covariances which enter in the definitions of  $\mathbf{M}$  and  $\mathbf{N}$  are related to the transition matrix of the Markov chain [1], and hence depend not only on the model, but also on the dynamics. Finally, it is worth noting that in order to correctly obtain the FSS limit, the coefficients  $\{\lambda_i\}$  must be the same for every size. They can be, however, chosen differently for each observable  $O$  considered.

**Results** We have tested this method on the standard Ising model, which is defined on a  $d$ -dimensional lattice with linear size  $L$  by the Hamiltonian

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \quad \sigma_i = \pm 1, \quad (14)$$

where the sum is over the nearest-neighbors sites and one can set  $J = 1$ . We consider here four RG-invariant quantities.

$L$	$\chi$	$(dU_4/d\beta)/L^d$	$(dU_6/d\beta)/L^d$	$(dR_\xi/d\beta)/L^d$
Metropolis dynamics				
<b>8</b>	88.19(53)	-0.1112(15)	-0.4758(80)	0.07330(49)
imp	88.90(15)	-0.11412(42)	-0.5341(21)	0.08078(35)
gain	11.7	12.3	14.2	2.0
<b>16</b>	351.0(3.3)	-0.04140(86)	-0.1813(48)	0.02652(29)
imp	354.91(63)	-0.04269(29)	-0.19932(98)	0.02854(13)
gain	27.9	9.1	24.2	4.9
<b>32</b>	1382(14)	-0.01574(36)	-0.0697(20)	0.00992(11)
imp	1389.3(2.6)	-0.015837(75)	-0.07383(40)	0.010367(46)
gain	26.9	23	25.8	6.2
<b>64</b>	5373(46)	-0.00612(12)	-0.02731(70)	0.003686(34)
imp	5435.9(9.8)	-0.005953(31)	-0.02760(14)	0.003822(17)
gain	22.2	16.1	25.2	4.2
<b>128</b>	21537(200)	-0.002220(48)	-0.00981(27)	0.001412(15)
imp	21141(40)	-0.002215(11)	-0.010304(56)	0.0014148(61)
gain	25.5	19.7	23.	6.1
Wolff dynamics				
<b>8</b>	87.82(28)	-0.1098(11)	-0.4682(55)	0.07244(40)
imp	88.48(11)	-0.11608(37)	-0.5418(20)	0.07995(25)
gain	6.3	8.9	7.7	2.7
<b>16</b>	350.7(1.2)	-0.04273(45)	-0.1880(23)	0.02704(16)
imp	352.83(47)	-0.04306(17)	-0.20094(90)	0.028609(91)
gain	7.0	7.0	6.7	3.1
<b>32</b>	1381.2(5.3)	-0.01604(18)	-0.07162(92)	0.009997(63)
imp	1388.4(1.9)	-0.016051(76)	-0.07521(39)	0.010391(36)
gain	7.8	5.6	5.5	3.1
<b>64</b>	5393(22)	-0.006101(75)	-0.02746(39)	0.003749(27)
imp	5438.6(7.3)	-0.005943(33)	-0.02769(16)	0.003824(15)
gain	9.5	5.3	6.3	3.2
<b>128</b>	21237(94)	-0.002289(30)	-0.01025(16)	0.001410(10)
imp	21188(29)	-0.002291(14)	-0.010308(64)	0.0014145(59)
gain	10.1	4.6	5.8	3.1

TABLE II. Same as in Table I for the 3D Ising model.

Besides the cumulants  $U_4$  and  $U_6$  defined by

$$U_{2j} \equiv \frac{\langle M^{2j} \rangle}{\langle M^2 \rangle^j}, \quad M \equiv \sum_i \sigma_i, \quad (15)$$

we also consider the ratio of the second-moment correlation length  $\xi$  over  $L$ ,  $R_\xi \equiv \xi/L$ , as well as the ratio of the partition function of a system with antiperiodic boundary conditions (b.c.) on one direction over the partition function of a system with full periodic b.c.  $R_Z \equiv Z_a/Z_p$ ; this quantity can be efficiently sampled using the boundary flip algorithm [11]. In Eq. (15) the brackets  $\langle \rangle$  denote the thermal average. As for the observables, we consider here the susceptibility  $\chi \equiv \langle M^2 \rangle / L^d$ , and the derivatives of RG-invariant quantities  $dU_4/d\beta$ ,  $dU_6/d\beta$ ,  $dR_\xi/d\beta$ . At the critical point, as well as in the FSS at fixed phenomenological coupling, one has  $\chi \propto L^{2-\eta}$  and  $dR/d\beta \propto L^{1/\nu}$ ,  $R = U_4, U_6, R_\xi$ , where  $\eta$  and  $\nu$  are critical exponents. Thus these observables

can be used to extract the value of the critical exponents. As mentioned after Eq. (13), the coefficients  $\{\lambda_i\}$  can be optimized separately for each observable. However, for a given observable they have to be the same for each lattice size. Of course, the coefficients determined from Eq. (11) will be in general different for each lattice. Since in a typical simulation most of the computational time is spent on the largest lattice, a possible strategy for choosing the coefficients  $\{\lambda_i\}$  consists in applying Eq. (11) to the largest available lattice, and then to use the same coefficients for the smaller lattices. We have found that the gain in the CPU time in most of the cases does not vary significantly with the lattice size, despite the fact that the chosen  $\{\lambda_i\}$  are not the optimal ones for the smaller lattices. We have simulated the Ising model for lattice sizes  $L = 8-128$  in two and three dimensions. As mentioned above, the statistical error bars and hence the gain in the CPU time resulting from the present method depend not only on

the model, but also on the dynamics used. To illustrate this aspect, we have considered data sampled using, separately, two different simulation algorithms: Metropolis and Wolff single cluster. Simulations in  $d = 2$  have been carried out at  $\beta = 0.4406867935$ . The analysis at fixed phenomenological coupling has been done using the critical-point values  $U_4 = 1.167923(5)$ ,  $U_6 = 1.455649(7)$ ,  $R_\xi = 0.9050488292(4)$  [12], and  $R_Z \simeq 0.3728848808$  [13]. The results are reported in Table I. We compare the observables as determined with a standard analysis and with the present method (imp), with the coefficients determined, for each observable, from the data at  $L = 128$ . Notice that the mean values are not expected to coincide in the two analyses, although they can be very close. The gain factor in the CPU time is given by the square of the ratio of the error bars. Inspecting Table I we observe large gains in the CPU time, especially for the observables  $\chi$ ,  $dU_4/d\beta$ ,  $dU_6/d\beta$  sampled with the Metropolis algorithm, where in most of the cases the gain lies between 20 and 50; for the same observables, gains roughly between 6 and 9 are obtained in the case of Wolff dynamics. The gain in the CPU time obtained for  $dR_\xi/d\beta$  is roughly  $\sim 2$ –3 for both algorithms. Simulations in  $d = 3$  have been carried out at  $\beta = 0.2216544$ . FSS analysis at fixed phenomenological coupling has been done using recent determinations of the critical-point values  $U_4 = 1.6036(1)$ ,  $U_6 = 3.1053(5)$ ,  $R_\xi = 0.6431(1)$ ,  $R_Z = 0.5425(1)$  [14]. The results are reported in Table II. The gains in the CPU time for  $\chi$ ,  $dU_4/d\beta$ ,  $dU_6/d\beta$  sampled with the Metropolis dynamics are in most of the cases between 20 and 30. For the other observables

and for the Wolff dynamics the gains are similar to the bidimensional case. The present method represents also a significant improvement over a simpler approach of fixing one of the four RG-invariant quantities: for instance by fixing  $R_\xi$  the CPU gain factor for  $\chi$  is  $\lesssim 4$  in  $d = 2$ ,  $\lesssim 9$  in  $d = 3$  with Metropolis dynamics, and  $\lesssim 6$  in  $d = 3$  with Wolff dynamics.

To summarize, we have presented a FSS method which allows for a substantial reduction of the statistical error bars without additional computational time. Application of the method to the Ising model shows that the CPU gain factor does not depend much on the lattice size, and it is more pronounced for a local Metropolis update algorithm. Other possible interesting applications of this method should be found in “improved models” [3], where leading scaling corrections are suppressed, and in models with quenched disorder [15], where cluster algorithms are available only in special cases and significant reduction of error bars with FSS at fixed  $R_\xi$  has been reported [8].

## ACKNOWLEDGMENTS

The author is grateful to Martin Hasenbusch for useful discussions and communications and to Roderich Moessner and Ettore Vicari for useful comments on the manuscript. The author acknowledges support from the Max-Planck-Institut für Metallforschung (Stuttgart, Germany), where part of the work has been done.

- 
- [1] A. Sokal, in *Functional Integration: Basics and Applications (1996 Cargèse Summer School)*, NATO Science Series B: Physics, Vol. 361, edited by C. Dewitt-Morette, P. Cartier, A. Folacci (Plenum Press, New York, 1997), p. 131
  - [2] D. P. Landau, K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, England, 2005)
  - [3] A. Pelissetto, E. Vicari, Phys. Rep. **368**, 549 (2002)
  - [4] M. Weigel, W. Janke, Phys. Rev. Lett. **102**, 100601 (2009); Phys. Rev. E **81**, 066701 (2010)
  - [5] L. A. Fernández, V. Martín-Mayor, Phys. Rev. E **79**, 051109 (2009)
  - [6] M. Hasenbusch, J. Phys. A **32**, 4851 (1999)
  - [7] M. Hasenbusch, A. Pelissetto, E. Vicari, J. Stat. Mech.: Theory Exp. P12002 (2005)
  - [8] M. Hasenbusch, F. Parisen Toldin, A. Pelissetto, E. Vicari, J. Stat. Mech.: Theory Exp. P02016 (2007)
  - [9] D. Amit, V. Martín-Mayor, *Field Theory: The Renormalization Group and Critical Phenomena*, 3rd ed. (World Scientific, Singapore, 2005)
  - [10] H. G. Ballesteros, L. A. Fernández, V. Martín-Mayor, A. Muñoz Sudupe, Nucl. Phys. **B 483**, 707 (1997)
  - [11] M. Hasenbusch, Physica A **197**, 423 (1993)
  - [12] J. Salas, A. Sokal, J. Stat. Phys. **98**, 551 (2000)
  - [13] M. Caselle, M. Hasenbusch, Nucl. Phys. **B 470**, 435 (1996)
  - [14] M. Hasenbusch, Phys. Rev. B **82**, 174433 (2010)
  - [15] N. Kawashima, H. Rieger, in *Frustrated Spin Systems*, edited by H. T. Diep (World Scientific, Singapore, 2004)